

Modeling spike occurrences in electricity spot prices for forecasting

Citation for published version (APA):

Eichler, M., Grothe, O., Manner, H., & Türk, D. D. T. (2012). *Modeling spike occurrences in electricity spot prices for forecasting*. METEOR, Maastricht University School of Business and Economics. METEOR Research Memorandum No. 029 <https://doi.org/10.26481/umamet.2012029>

Document status and date:

Published: 01/01/2012

DOI:

[10.26481/umamet.2012029](https://doi.org/10.26481/umamet.2012029)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Michael Eichler, Oliver Grothe, Hans
Manner, Dennis Tuerk

**Modeling spike occurrences in
electricity spot prices for
forecasting**

RM/12/029



METEOR

Maastricht University School of Business and Economics
Maastricht Research School of Economics
of Technology and Organization

P.O. Box 616
NL - 6200 MD Maastricht
The Netherlands

Modeling spike occurrences in electricity spot prices for forecasting

Michael Eichler^a, Oliver Grothe^b, Hans Manner^b, Dennis Tuerk^{a,*}

^a*Department of Quantitative Economics, Maastricht University, The Netherlands*

^b*Seminar für Wirtschafts- und Sozialstatistik, Universität zu Köln, Germany*

Abstract

Predicting the occurrence of extreme prices, so-called spikes, is one of the greatest challenges when modeling electricity spot prices. Despite the fact that recently new insights have been achieved, the contemporaneous literature seems to be still at its beginning of understanding the different mechanisms that drive spike probabilities. We therefore reconsider the problem of forecasting the occurrence of spikes, in the Australian electricity market. For this purpose, we first discuss properties of the price data with a focus on the occurrence of spikes. We then propose simple models for the probability of spikes which take these properties into account. The models compare favorably for in- and out-of-sample forecasts to a competing approach based on the autoregressive conditional hazard model.

Keywords: electricity spot prices, price spikes, forecasting binary variables, autoregressive conditional hazard, dynamic logit

1. Introduction

One of the most important stylized features of electricity spot prices—besides mean reversion and strong seasonality—is the infrequent occurrence of spikes. These are periods of extreme prices that are typically short-lived and during which the spot price exceeds its normal level many times over. As pointed out by, e.g., Knittel and Roberts (2005), spikes occur due to price inelastic market participants when either the demand varies (often as a result of weather conditions) or the supply decreases, e.g., due to outages of generators or transmission lines. The effects can be amplified by generator bidding behavior, to be described later, as the central dispatch process does sometimes require to call generators of high marginal costs into production to satisfy demand. Thomas et al. (2011) state that extreme spikes are even more prevalent in markets with compulsory participation, as is the case in Australia's National Electricity Market (NEM).

Price spikes constitute a major source of risk for market participants such as retailers, who buy electricity for a flexible price from the NEM but sell for a fixed price to consumers. Advanced knowledge about the occurrence of spikes can help to reduce this risk. For instance, retailers could incur contracts with industrial consumers that allow them to interrupt supply for a certain amount of hours a day whenever prices exceed a specified threshold. Furthermore they could close bilateral contracts with producers, paying a premium for producers to insure them against extreme prices. For this purpose it would be useful to know when prices are prone to spike. At the same time producers could make use of this knowledge to set strategic bids. Moreover, industries that are able to shift energy consumption over short periods of time such as refrigerated warehouses or aluminium smelters could make use of spike forecasts to reduce spending on electricity.

The relevance of modeling price spikes has been acknowledged in the literature, and former studies (see Mount et al., 2006, Kanamura and Ohashi, 2008 and Christensen et al., 2009) have already successfully treated the subject of time-varying transition probabilities between extreme and moderately price regimes. In this context Christensen et al. (2012) recently proposed the

*Corresponding author

Email address: d.tuerk@maastrichtuniversity.nl (Dennis Tuerk)

autoregressive conditional hazard (ACH) model to forecast the occurrence of price spikes in the Australian electricity market on a half hourly basis. To our best knowledge they are the first to address the subject of intraday real time forecasting of extreme electricity spot prices.

The ACH model as applied by Christensen et al. (2012) is based on the assumption that the rate of spike occurrence between two spikes depends only on exogenous factors such as loads and temperature but is otherwise constant, where a spike is defined as an extreme spot price exceeding some chosen threshold. Although the ACH model improves upon a simpler model in which the spike occurrence only depends on the exogenous variables (see Christensen et al., 2012), it is not capable of capturing some important features of spike occurrences in electricity spot prices in the Australian markets. In particular, we found that periods of extreme prices often last one hour and longer and tend to occur again after 24 hours. Moreover, such periods are mostly preceded by periods of increasing prices. In order to capture these features, we consider two extensions of the original ACH model by Christensen et al. (2012) that include additional information about the past of the price process. Furthermore, we discuss four dynamic logistic regression models as alternatives to the ACH specifications. The simplest of these models has a hazard rate that depends only on the occurrence of a spike in the previous period and serves as a baseline model for the logit specification. The remaining models are inspired by the work of Kauppi and Saikkonen (2008) and incorporate additional past information and exogenous variables. The proposed models yield superior in-sample model fits as well as improved out-of-sample forecast properties.

The rest of the paper is structured as follows. In Section 2 we describe the Australian electricity market and analyze the features of the data. In Section 3 we present the statistical models studied in this paper. Additionally to models for the occurrence of spikes, we briefly discuss forecasting electricity loads, which are known to drive a large part of the observed price movements and are required by all but one of the considered models. In particular, we propose a novel forecasting scheme for loads that performs favorably compared to a standard approach proposed by Weron (2006) and used by Christensen et al. (2012). Section 4 gives a thorough comparative forecast evaluation of the models and a conclusion can be found in Section 5.

2. Market structure and data description

In this section, we provide a detailed analysis of the occurrence patterns of price spikes. While most models for electricity spot prices assume that spikes occur more or less randomly over time, our results show that there are clear dependence patterns, which are neglected in these traditional approaches. Similar observations can be found in Christensen et al. (2009, 2012). We start our empirical analysis by a brief description of the structure of the Australian electricity market. More extensive details can be found in AEMO (2010), Chan et al. (2008) or Anderson et al. (2007).

2.1. The Australian electricity market

Since December 1998 the Australian National Electricity Market (NEM) is operating as a wholesale market supplying electricity to retailers and end-users in Queensland (QLD), New South Wales (NSW), Victoria (VIC) and South Australia (SA). Tasmania joined the NEM in 2006. The Snowy region located in southern NSW was initially a separate market in the NEM, but was abolished on the 1st of July 2008 and its generation assets were incorporated into the NSW and VIC regions. Thus, today operations are based in five interconnected regions that largely follow state boundaries. The Australian Electricity Market Operator (AEMO) facilitates the needed infrastructure, allowing for more than 10 billion Australian dollar of electricity to be traded per year.

Exchange between producers and retailers/consumers is facilitated through a pool in which output of all generators is aggregated and scheduled to meet forecasted demand. In contrast to most other electricity markets the wholesale trading is conducted as a real-time market where supply and demand are instantaneously matched through a centrally coordinated dispatch process. Each of the approximately 260 registered generators is allowed to submit up to ten price-volume combinations for each 5 minute interval of the next day (daily bids) before 12:30 pm. However, as opposed to pure day-ahead markets, the generator has the right to change the corresponding volume for each submitted price (re-bid) up to 5 minutes before execution. This is almost equivalent to a

Table 1: Descriptive statistics for extreme prices above A\$100/MWh and above A\$300/MWh.

	VIC		NSW		QLD		SA	
Threshold	A\$100	A\$300	A\$100	A\$300	A\$100	A\$300	A\$100	A\$300
Mean	452.0	2069.9	623.1	2140.1	564.5	2163.0	622.4	3169.0
Median	145.9	1070.5	163.4	863.1	164.5	1352.2	147.9	1592.9
Std dev	1191.5	2454.9	1509.3	2591.2	1283.1	2242.5	1786.2	3596.7
Skewness	5.65	1.91	4.21	1.62	4.26	1.52	4.55	1.14
Kurtosis	38.45	5.93	21.66	4.70	22.69	4.53	23.34	2.71
Number of spikes	2727	424	3347	784	3250	650	3979	616
Percentage spikes	1.56	0.24	1.91	0.45	1.85	0.37	2.27	0.35

continuous trading setup. From all offers submitted, AEMO determines the generators required to produce electricity based on the principle of meeting prevailing demand in the most cost-efficient way and dispatches them into production. In accordance with the bidding structure such a dispatch price is determined every 5 minutes and represents the cost to supply the last megawatt of electricity to meet demand. The resulting price applies to all generators scheduled into production regardless of the level of their original offer. Six dispatch prices are averaged every half-hour to determine the spot price, yielding 48 trading intervals for each day and region. These spot prices are used as the basis for the settlement of financial transactions for all energy products traded in the NEM and are publicly available from AEMO.

Extreme positive price spikes can occur in case that demand exceeds moderate priced supply and offers of high pricing generators apply. This phenomena usually occurs through unexpected increases in demand and when generators or transmission lines fall out (see, e.g., the discussion about price formation in electricity markets in Knittel and Roberts (2005) and the references therein). Given the quasi-continuous trading in the Australian market, intraday forecasts of price spike probabilities are especially relevant for market participants.

2.2. Data-analysis

Our data set consists of half-hourly spot prices, i.e., the highest frequency available, for the four main Australian markets Victoria (VIC), New South Wales (NSW), Queensland (QLD) and South Australia (SA) starting 1 January 2002 and ending on 31 December 2011. Data from earlier periods were omitted from the analysis because the four markets became physically interconnected only in the year 2001.

For our analysis, we distinguish between moderate prices, price spikes, and extreme price spikes. Following Christensen et al. (2012), we define prices exceeding a threshold of A\$100/MWh as spikes whereas prices above A\$300/MWh are considered as extreme spikes. Although these choices appear somewhat arbitrary, the lower threshold of A\$100/MWh seems to be widely accepted by market participants (see Christensen et al., 2012) while the higher threshold corresponds to the strike price of heavily traded derivatives. Note that for convenience we restrict the graphical data analysis to spikes defined as prices greater A\$100/MWh while reporting the modeling results for both thresholds.

Negative prices can potentially occur with a floor of -A\$1 000/MWh but are very rare, whereas positive prices are capped at a maximum of A\$12 500/MWh.¹ Descriptive statistics concerning price spikes are given in Table 1. The positive differences between mean and median as well as the positive skewness indicate that for all markets price spikes are right skewed. The high standard deviation and kurtosis show how erratic prices behave when spiking. Furthermore the proportion of spikes varies between 1.6% and 2.4% of all prices while 0.23% and 0.44% when defining spikes as prices greater A\$300/MWh. Figure 1 shows a 'barcode' plot depicting the occurrence of spikes for the year 2010 for all markets.² It can be seen from the graph that spikes do not only occur in

¹The maximum price cap was increased from A\$5 000/MWh to A\$10 000/MWh in April 2002 and again from A\$10 000/MWh to A\$12 500/MWh in July 2010.

²For other time periods the graphs look very similar and are available upon request.

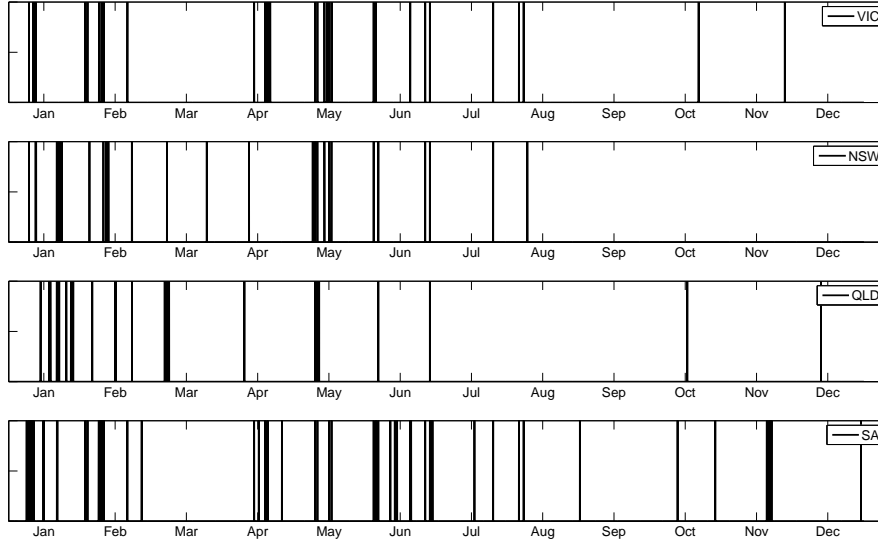


Figure 1: Spike occurrences during the year 2010.

clusters, but that in fact spikes occur in blocks of varying length. Figure 2 shows histograms of the block length for all four markets for the entire data set. The histogram shows that a characteristic property of spikes is that once a spike has occurred it is very likely to be followed by further spikes. For daily data such behavior is considered a stylized fact of electricity prices (Janczura and Weron, 2010). For intraday data, however, this has not yet been explicitly acknowledged.

Finally, Figure 3 shows the probabilities of spike occurrences conditional on a previous spike for each of the four electricity markets. One can see a clear pattern that is identical for all four markets. The pattern is in accordance with our previous finding that spikes are often followed by further spikes. Furthermore, spikes are likely to be followed by spikes at the same time on the following day. One may suspect that this is due to the fact that spikes always occur at the same time of the day when demand is high. However, Figure 4 showing the distribution of spikes and loads over all half-hour intervals of the week suggests that this is only partly true. Although most spikes occur close to midday and the early evening, one can see that spikes are distributed over most parts of the day with the exception of night hours. Furthermore, the seasonality of the (standardized) loads resembles the behavior of the spikes closely.

We follow the choices of previous studies (see e.g. Rambharat et al., 2005, Mount et al., 2006 and Christensen et al., 2012) and use (forecasts of) loads and temperatures as exogenous variables. Loads are available from AEMO on a half hourly basis, whereas data of temperatures can be obtained from the Australian Bureau of Meteorology on a daily basis. Including these variables allows to capture a large part of the seasonal patterns in spike occurrence. This holds for daily and weekly seasonality (loads), as well as seasonality over the year (temperatures). Consequently, directly modeling the seasonality is not necessary and does not lead to a better model performance. Note that for the weather data the same temperature is assumed for each half hour interval during a day and we use actual temperatures instead of forecasts. We believe that this simplification is innocuous because weather forecasts of high quality are available from meteorological services.

Christensen et al. (2012) state that loads exhibit a trend in mean and volatility and for this reason should be adequately standardized. For each observation they suggest to subtract the mean and divide by the standard deviation of the previous 365 days. To further correct for the positive correlation between temperature and prices in summer and the reverse occurrence in winter, the authors propose to use the daily absolute deviations of the minimum and maximum temperatures from their averages, which are again calculated over the preceding 365 days. In the following we will refer to these corrected values when referring to loads, minimum and maximum temperatures

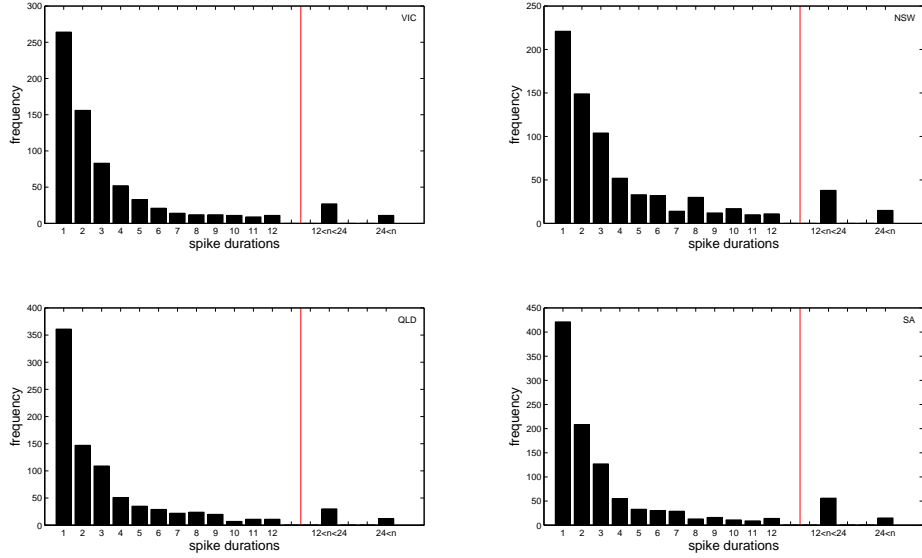


Figure 2: Histograms of the durations of spikes for VIC, NSW, QLD, and SA. The duration of a spike is given by the number of consecutive spike events.

and denote them by L_t , $W_{min,t}$, and $W_{max,t}$, respectively.

3. Models to forecast spikes

In this section we deal with the problem of modeling and forecasting spikes in half-hourly electricity spot prices. The models we consider use loads as an exogenous variable. Therefore, we first have to address the problem of forecasting loads, which is done in Section 3.1. We compare our approach to an existing method for forecasting loads.

To our knowledge Christensen et al. (2012) were the first to address the problem of spike forecasting in a high frequency context. For this purpose the authors used the autoregressive conditional hazard (ACH) model, which we present in Section 3.2 together with two extensions we suggest to account for the specific features of the spike series discussed earlier. As an alternative to the ACH approach in Section 3.3, we also consider dynamic logistic regression models that incorporate past information concerning prices at specific lags.

3.1. Forecasting loads

Most model specifications that we consider in this paper include the electricity loads as exogenous variable since these are known to drive the movements in the electricity spot price to a large extent. When using these models for forecasting the occurrence of spikes, we therefore require reliable forecasts of the loads themselves. Apart from this, short-term load forecasts are of independent interest as they form the basis to scheduling generation and transmission of electricity (e.g., Taylor et al., 2006). Consequently, a wide range of literature concerning short-term load forecasting has emerged during the last years. Distinct contributions to load forecasting based on univariate methods were made by Taylor (2003) and more recently by Taylor (2012), who applied different techniques that were presented in Taylor (2010). Furthermore Fan and Hyndman (2012) propose a new statistical methodology to forecast the short-term demand for two regions of the Australian electricity market including exogenous variables.

Despite its general importance load forecasting will not be the focus of the present paper. Therefore, we implemented a simple univariate approach that does not rely on numerical optimization methods. We note that loads considered in this section are standardized as described in Section 2.2.

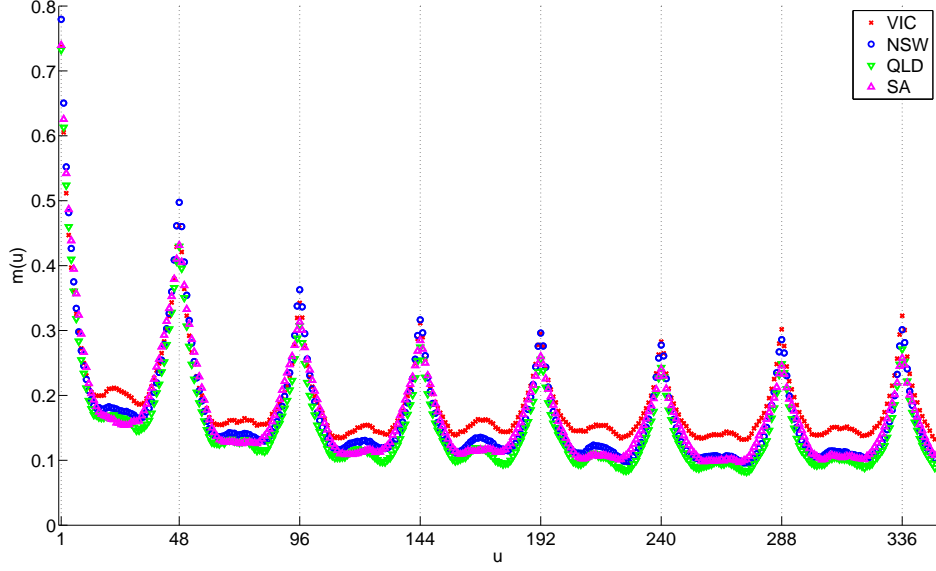


Figure 3: Correlogram of spike occurrences for the four markets VIC, NSW, QLD, and SA. The correlogram gives the conditional intensity $m(u)$ for a spike at time u given that a spike has occurred at time $t = 0$.

Loads exhibit a very stable periodic intra-weekly pattern. Our forecasting approach combines knowledge of this weekly seasonality with exponentially weighted moving averages to get daily and half-hourly load forecasts. Let J be the total number of weeks and let $L_{i,j}$ with $i \in 1, \dots, 336$ and $j \in 1, \dots, J$ be the i -th half-hourly load in week j out of the total of $48 \cdot 7 = 336$ half hours of the week. Further let

$$\bar{L}_i = \frac{1}{J} \sum_j L_{i,j}, \quad i \in 1, \dots, 336, \quad j \in 1, \dots, J,$$

be the mean of the loads at the i th weekly half-hour shown in panel 2 of Figure 4.

We use these estimates to further deseasonalize the corresponding observations of the loads, yielding $l_t = L_{i,j} - \bar{L}_i$ with $t = i + (j - 1)336$, which can be interpreted as deseasonalized loads at half hour t with respect to the weekly seasonality. We denote a forecast of the load at time t based on information up to time $t - 48$, i.e., a daily forecast, by \hat{L}_t and use

$$\hat{L}_t = \bar{L}_i + \sum_{k=1}^7 \omega_k l_{t-48k}$$

with weights

$$\omega_k = \frac{\exp(-k)}{\sum_{m=1}^7 \exp(-m)}.$$

Thus, the daily forecast is done by applying exponentially decaying weights to the deseasonalized observations that happened at the same time of the day over the past week and re-adding the value of the intra-week pattern for time t .

Similarly, we denote forecasts for time t based on information up to time $(t - 1)$, i.e., half hourly forecasts, by \hat{L}_t^* . We propose to use the daily forecast for time t corrected by the forecast error from half an hour earlier,

$$\hat{L}_t^* = \hat{L}_t + (L_{t-1} - \hat{L}_{t-1}).$$

To evaluate these forecasting approaches we apply them to load data from 1st of January 2002 to 31st of December 2012 and compare the results to the ones yielded when using a naive method, that was applied in Christensen et al. (2012). This simple approach forecasts loads for Saturdays,

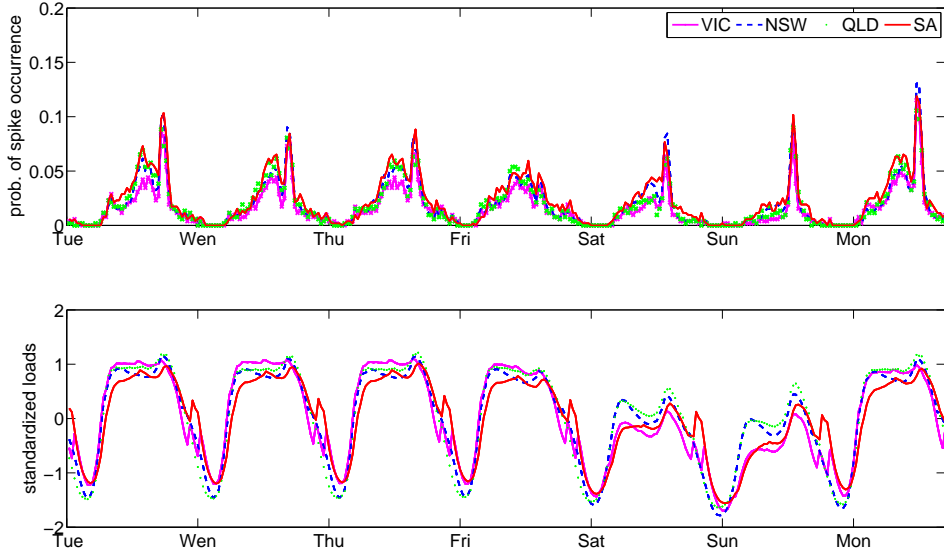


Figure 4: Weekly seasonality of spikes (top) and loads (bottom)

Table 2: Loss-functions to evaluate forecast errors of loads

	VIC		NSW		QLD		SA	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
\hat{L}_{t+1}^W	0.3101	0.5071	0.2508	0.3940	0.1994	0.3131	0.4290	0.6844
\hat{L}_{t+1}	0.2591	0.4094	0.1980	0.3063	0.1567	0.2407	0.3435	0.5185
\hat{L}_{t+1}^*	0.0505	0.0709	0.0414	0.0572	0.0386	0.0531	0.0610	0.0816

Sundays and Mondays at each point in time by the loads recorded for the same time and day of the previous week and was suggested by Weron (2006) as a benchmark for more sophisticated models. Table 2 shows the mean absolute error (MAE) and the root mean squared error (RMSE) for our forecasting approach and the naive method, which is denoted by \hat{L}_{t+1}^W .

The results show that the daily forecasts with the proposed model are consistently superior to those yielded with the naive method that was described in Weron (2006), giving smaller loss functions for every market. Our half hourly forecasts lead to loss functions that are approximately 5 times smaller than when using our daily forecasting approach.

3.2. Autoregressive conditional hazard models

We now turn to the models for the occurrence of price spikes. These occurrences can be modelled by the binary process S_t that takes value 1 whenever the spot price exceeds the specified threshold (here A\$100/MWh or A\$300/MWh respectively) and zero otherwise. Such a binary process is characterized by its hazard rate, that is, by the probability of a spike occurring at time t conditional on the past history up to time $t - 1$. Thus the hazard rate is given by

$$h_t = \text{Prob}(S_t = 1 | \mathcal{H}_{t-1}),$$

where \mathcal{H}_{t-1} denotes all relevant information observed up to time $t - 1$. In the autoregressive conditional hazard (ACH) model originally proposed by Hamilton and Jorda (2002), the hazard rate depends only on the observed durations between (spike) events. Since the hazard rate is thus constant between spike occurrences, the durations are geometrically distributed. More precisely, denoting the duration between the $(n - 1)$ th and n th spike by u_n and letting ψ_n be the expectation

of u_n given the previous durations u_{n-1}, u_{n-2}, \dots , the hazard rate h_t at time t is given by

$$h_t = \frac{1}{\psi_{N(t)+1}}, \quad (3.1)$$

where $N(t)$ is the counting process giving the total number of spike events up to time t . The main idea of the ACH model and the underlying autoregressive conditional duration (ACD) model by Engle and Russel (1998) is that the intensities ψ_n depend only on the past durations and intensities similar to a GARCH specification by

$$\psi_n = \omega + \sum_{j=1}^k \alpha_j u_{n-j} + \sum_{j=1}^l \beta_j \psi_{n-j}. \quad (3.2)$$

As the distribution of the durations is extremely skewed, Christensen et al. (2012) apply a generalized version of the ACH model proposed by Fernandes and Grammig (2006), which makes use of a Box-Cox transformation to reduce the skewness of the durations. Thus the intensities ψ_n satisfy

$$\psi_n^\nu = \omega + \sum_{j=1}^k \alpha_j u_{n-j}^\nu + \sum_{j=1}^l \beta_j \psi_{n-j}^\nu, \quad (3.3)$$

where the parameter ν controls the Box-Cox transformation. The original ACH model is obtained for $\nu = 1$.

The use of calendar time in (3.1) allows to include additional explanatory variables, such as loads and temperatures, which typically vary between spike occurrences in the model. For instance, Christensen et al. (2012) consider the modified hazard rate

$$h_t = \frac{1}{\Lambda(\psi_{N(t-1)} + \exp(-\gamma' z_t))}, \quad (3.4)$$

where $z_t = (z_{1,t}, \dots, z_{k,t})'$ is the vector of explanatory variables. The function Λ was proposed by Hamilton and Jorda (2002)³ to ensure that h_t satisfies the constraint $0 < h_t < 1$. The authors further propose to omit the intercept in (3.3) and add a constant to z_t in (3.4). The parameters of the model are estimated by maximization of the likelihood function.

In the model used by Christensen et al. (2012), the regressor z_t consisted of a constant and the (forecasted) load \hat{L}_t as well as the temperatures $W_{min,t}$ and $W_{max,t}$ as exogenous variables. For the linear regressor part in (3.4) this yields

$$\gamma' z_t = \gamma_0 + \gamma_1 \hat{L}_t + \gamma_2 W_{min,t} + \gamma_3 W_{max,t}. \quad (3.5)$$

In the remainder of the paper, we will refer to this model as *ach1* model. Motivated by the characteristics of the data we extend the *ach1* model and propose two further specifications within the ACH framework.⁴ The first extension includes the log-prices, p , at times $(t-1)$ and $(t-48)$ and is based on the findings from Section 2. The resulting specification for the conditional spike probabilities is then

$$\gamma' z_t = \gamma_0 + \gamma_1 \hat{L}_t + \gamma_2 W_{min,t} + \gamma_3 W_{max,t} + \gamma_4 p_{t-1} + \gamma_5 p_{t-48}, \quad (3.6)$$

and thus the same as the one for the *ach1* model with two more exogenous variables, p_{t-1} and p_{t-48} . We will call this specification *ach2* model in the following.

To account for the block structure of spikes found earlier we suggest a further specification, called *ach3* model, by adding an interaction term that allows the effect of p_{t-1} and p_{t-48} to differ if a spike has occurred at $(t-1)$:

$$\gamma' z_t = \gamma_0 + \gamma_1 \hat{L}_t + \gamma_2 W_{min,t} + \gamma_3 W_{max,t} + \gamma_4 p_{t-1} + \gamma_5 p_{t-48} + (\gamma_6 p_{t-1} + \gamma_7 p_{t-48}) S_{t-1}. \quad (3.7)$$

³The precise form of the transformation is given in footnote 3 of Hamilton and Jorda (2002).

⁴As the equations would get very long we will not write the following two extensions down completely, but give their representations of $\gamma' z_t$ instead.

As defined earlier S_{t-1} is an indicator function that takes on the value one when a spike has taken place at time $(t - 1)$ and zero otherwise. We thus estimate a threshold autoregressive model that nests *ach2*, adding two parameters for variables p_{t-1} and p_{t-48} whenever $S_{t-1} = 1$.

Since the ACH models are able to capture the clustering of events, they appear to be good choices for the problem at hand. However, recall that electricity price spikes do not only appear in clusters, but usually in blocks of consecutive spikes. Although the ACH models include exogenous information that allows to account for this features, it is not obvious that the dynamics implied by the model are entirely appropriate. The reason is the fact that—without accounting for the exogenous information—the model might need to much time to adapt whenever a block of spikes starts or ends. In the following we propose simpler alternative models that take the features presented in our data description into account.

3.3. Dynamic logit models

In Section 2.2 we observed that the majority of spikes occur in blocks of more than one observation. A natural approach to model the occurrence of such clustered blocks is a logit model with past prices and exogenous information. As before, define h_t as the probability of a spike at time t given the information available up to time $(t - 1)$. Using a logistic regression, in this section, we then model h_t as

$$h_t = \frac{1}{1 + \exp(-\beta'x_t)}, \quad (3.8)$$

with $\beta'x_t$ defining the different model specifications.

As a first logit model, called *logit1* model, we propose a simple forecasting method that solely uses the fact that spikes do tend to occur in blocks:

$$h_t = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 S_{t-1}))}, \quad (3.9)$$

where S_{t-1} again is a spike indicator for period $(t - 1)$. With this logit model we obtain probability forecasts that correspond to the unconditional probability of a spike following a spike, $P(S_t = 1|S_{t-1} = 1)$ and the unconditional probability of a spike occurring after a moderate price, $P(S_t = 1|S_{t-1} = 0)$. Despite the fact that this approach immediately adapts after entering or leaving a spike block, it also has its disadvantages. In case of $P(S_t = 1|S_{t-1} = 1) \geq 0.5$ and $P(S_t = 1|S_{t-1} = 0) < 0.5$ one will, by construction, never be able to forecast the first spike of a block and will make one false prediction whenever a spike-block ends if a probability of 0.5 is taken as basis for this decision. As for periods in which spikes tend to occur in long blocks, this model still should be quite appropriate, we decide to use it as a benchmark for further model specifications.⁵

Next we define a counterpart to the *ach2* model, including the same exogenous variables:

$$\beta'x_t = \beta_0 + \beta_1 \hat{L}_t + \beta_2 W_{min,t} + \beta_3 W_{max,t} + \beta_4 p_{t-1} + \beta_5 p_{t-48} \quad (3.10)$$

which we call the *logit2* model. Again, β_0 does correspond to the constant term of the model while \hat{L}_t , $W_{min,t}$, and $W_{max,t}$ denote the load forecast for period t , and minimum and maximum temperature (defined in subsection 2.2) of the respective day. Further, as already done for the *ach2* model, we add lagged log prices, p_{t-1} and p_{t-48} . The latter variables capture the feature of spike blocks in a similar way as the *logit1* model with the difference that state-continuous lagged log prices are used instead of the binary spike indicator. The *logit2* model is thus capable of capturing sequentially occurring spikes, while allowing for further information to indicate spikes that have not been preceded by another spike and to announce the end of a spike block before it is actually over.

As an analogue to the *ach3* model, we again extend the *logit2* model adding an interaction term that allows the effect of p_{t-1} and p_{t-48} to differ if a spike has occurred at $(t - 1)$:

$$\beta'x_t = \beta_0 + \beta_1 \hat{L}_t + \beta_2 W_{min,t} + \beta_3 W_{max,t} + \beta_4 p_{t-1} + \beta_5 p_{t-48} + (\beta_6 p_{t-1} + \beta_7 p_{t-48}) S_{t-1}. \quad (3.11)$$

⁵As done already in context of the ACH specifications we will restrain ourselves to write down the equations concerning $\beta'x_t$ for further models, in order not to exceed one row for an equation.

Table 3: Models under investigation

	ach_1	ach_2	ach_3	$logit_1$	$logit_2$	$logit_3$	$logit_4$
Constant	×	×	×	×, S	×	×	×, B
$x_{N(t)-1}$	×	×	×	—	—	—	—
$\psi_{N(t)-1}$	×	×	×	—	—	—	—
L_t	×	×	×	—	×	×	B
$W_{min,t}$	×	×	×	—	×	×	B
$W_{max,t}$	×	×	×	—	×	×	B
p_{t-1}	—	×	×, S	—	×	×, S	×, B
p_{t-48}	—	×	×, S	—	×	×, S	×, B

Note: Parameter that are estimated for the whole time series are indicated with '×'. Parameters that are solely estimated for observations that follow prices greater than the threshold are indicated with an ' S ' and with a ' B ' if prices need to follow an observation that is smaller than the threshold. In case a parameter is not estimated at all in a certain model, we use '—' to indicate the fact.

The resulting *logit3* model thus nests the *logit2* model and can be interpreted as a dynamic threshold logit model that allows the effect of past log prices on h_t to depend on them exceeding a threshold in $(t-1)$.

A consequent extension of the idea that motivated us to construct the above model is to explicitly acknowledge the fact that price spikes are fundamentally different from moderate prices. This is done typically when modeling daily energy prices by distinguishing spike regimes and non-spike regimes. To this end we propose another simple regime switching model, called *logit4*. The model is defined as

$$\beta'x_t = c_1 + \beta_1 p_{t-1} + \beta_2 p_{t-48} + (1 - S_{t-1})(c_2 + \beta_3 \hat{L}_t + \beta_4 W_{min,t} + \beta_5 W_{max,t} + \beta_6 p_{t-1} + \beta_7 p_{t-48}). \quad (3.12)$$

Variable c_1 corresponds to the constant term when a spike is observed in $(t-1)$, while $c_1 + c_2$ yields the constant term when no spike is observed at $(t-1)$, in which case the indicator function $(1 - S_{t-1})$ switches from zero to one. The model further allows to isolate the effect of exogenous variables as load forecasts and temperatures to exclusively explain spike probabilities when observing moderate prices. If in contrast a spike has occurred at $(t-1)$, this exogenous information is disregarded and only past log-prices are assumed to be driving the probability of spikes. Thus we expect to allow higher explanatory power of exogenous variables when no spikes are prevailing.

Note that the ideas for the different logit specifications are similar to the ones proposed in Kauppi and Saikkonen (2008) in the context of U.S. recessions. The authors further propose to include h_{t-1} as an exogenous variable when forecasting h_t . This gives the logit model an appearance that is comparable to the ACH framework. In order to check the explaining power of this variable we also estimated models *logit2* to *logit4* including h_{t-1} as one of the exogenous variables. As results were rather mixed as well in- as out of sample, when comparing them to the original logit specifications presented earlier, we decided to discard h_{t-1} from the exogenous variables.

In order to allow the reader an easier overview of the different models and the parameters therein, we include Table 3. Each column corresponds to a different model. The use of variables in a certain model is indicated through '×' at the column that corresponds to the model and the rows that correspond to the variables. If variables are only to exercise an effect in case that a spike has occurred in the previous time interval, this is indicated by and ' S '. A ' B ' is used to denote that the effect is solely exercised when the previous observation has been of moderate nature.

4. Forecast comparison

In this section we report the results of our empirical analysis. We consider spikes defined as prices exceeding both A\$100/MWh and A\$300/MWh. Furthermore we use two different sample periods. The first one was chosen in order to compare our results to the ones of Christensen et al. (2012), so we chose 1st of January 2002 until 30th of September 2007. The out-of-sample period for

this sample is 1st of July until 30th of September 2007.⁶ We feel that this period is not entirely representative for the data set in the sense that very many spikes occur in the out-of-sample period compared to the whole dataset and that the most recent data is disregarded. Therefore we also report results for a second time window covering the period from 1st of January 2002 until 31st of December 2011, where the last year is considered for the out-of-sample evaluation. In Section 4.1 we report the in-sample fit for our models, whereas the out-of-sample results are reported in Section 4.2.

4.1. In-sample fit

We estimated all models presented in Section 3 with price spikes corresponding to thresholds of A\$100/MWh and A\$300/MWh for the two sample periods described above. In order to preserve space we do not report the parameter estimates here, but note that they are what is to be expected considering the data description in Section 2.2 and the estimates reported in Christensen et al. (2012).⁷ As the models are partly non-nested we decided to compare the in-sample fit with help of the Bayesian information criterion (BIC). Given any two estimated models, the one with the lower BIC value is the one to be preferred. Compared to other information criteria as, e.g., the Akaike information criterion it penalizes additional parameters quite strongly, so it should favor parsimonious specifications. As we are ultimately interested in models with a good forecasting performance we expect less heavily parameterized models to be more suitable. Based on the BIC we also computed the model confidence set (MCS) of Hansen et al. (2011). The MCS is a set of models whose performance is not significantly different considering a certain goodness-of-fit criterion and it can be seen as an analogue to a confidence interval for competing (non-nested) models. Thus we acknowledge that it is unlikely that a single model outperforms all the others, but that there are multiple models that perform equally well. The MCS is based on a sequence of hypothesis tests. It eliminates inferior models based on the criterion of interest. P-values for the sequential tests are determined by a block-bootstrap procedure as described in Hansen et al. (2011) and references therein. We used a size of 5%, 5000 bootstrap samples and a block length corresponding to 4 weeks of data. Table 4 reports the BIC for all estimated models, with bold numbers indicating that the MCS contains the corresponding model.

Looking at the results it stands out that the *logit4* specification always ranks first for both thresholds and sample periods. For a threshold of A\$100/MWh the MCS only contains other models for QLD. For a threshold of A\$300/MWh other models are included for all markets except NSW. Thus in terms of in-sample fit the rich parametrization of the *logit4* model appears to dominate all other models. Furthermore, looking at Table 4 it stands out that the *ach1* always performs far worse than all the other models.

4.2. Out-of-sample evaluation

In order to compare the forecasting performance of our models we estimated each using the in-sample period defined above and performed 1-step ahead forecasts of the probability to observe a spike. Whenever this forecast exceeded the value of 0.5 a spike was forecasted. Note that we did not re-estimate the model parameters each period, but that we only updated the information set. As a first way to compare the performance in Table 5 we report the number of detected spikes, as well as the correct detection rate (CDR) and the false detection rate (FDR). The CDR gives the ratio between correctly detected and observed spikes. The FDR comprises the ratio of falsely detected and the total number of detected spikes. Both, CDR and FDR are reported in percent. Bold numbers indicate which model gave the smallest value for the FDR and the largest value in case of the CDR. As to be expected, there is a tradeoff between correct and false detection, so it is not surprising that no single model outperforms the others in both measures. Nonetheless

⁶Note that despite the identical sample period the results reported for the *ach1* model will somewhat differ from the ones reported by the aforementioned authors. Reason is the fact that we use load forecasts instead of actual loads.

⁷Detailed estimation results are available from the authors upon request.

Table 4: Bayesian information criterion

	VIC		NSW		QLD		SA	
Threshold	A\$100	A\$300	A\$100	A\$300	A\$100	A\$300	A\$100	A\$300
Window I								
<i>ach</i> ₁	11629	2605	14671	3822	14672	4386	15692	2984
<i>ach</i> ₂	5638	1684	5561	2122	6577	2647	9103	2234
<i>ach</i> ₃	5552	1662	5701	2313	7102	2632	9135	2397
<i>logit</i> ₁	7458	2057	8544	3062	8850	3472	11505	2814
<i>logit</i> ₂	5434	1411	5712	1954	6594	2431	9118	2174
<i>logit</i> ₃	5206	1414	5551	1907	6294	2441	8736	2177
<i>logit</i> ₄	5032	1411	5273	1816	5984	2368	8473	2165
Window II								
<i>ach</i> ₁	18197	3807	19747	5149	22152	6186	23592	4614
<i>ach</i> ₂	9607	2281	8496	3048	10011	3926	13066	3689
<i>ach</i> ₃	8981	2452	8518	3135	9967	4041	13618	3657
<i>logit</i> ₁	12055	3041	12432	4185	13863	5051	16391	4438
<i>logit</i> ₂	9344	2174	8707	2834	11050	3704	13656	3367
<i>logit</i> ₃	8695	2160	8335	2762	10238	3700	12698	3350
<i>logit</i> ₄	8417	2112	7914	2590	9637	3615	12025	3275

Note: The table contains the BIC for the 7 different model specifications with spike thresholds (thr) equal to A\$100/MWh and A\$300/MWh and time-periods 1st of January 2002 to 30th of June 2007 (window I), as well as 1st of January 2002 to 31st of December 2010 (window II). Values that correspond to models belonging to the model confidence set (MCS) with a size of 5% are printed in bold. To calculate the MCS a block-bootstrap procedure with 336*4 observations per block and a total of 5 000 simulations was used.

it stands out that in most cases the *ach1* model shows the worst performance in terms of both correct *and* false detections. The logit models, in particular the *logit1* and *logit4* model, show the best performance with respect to correct detections. In terms of false detections no single model dominates, although *logit3* and *ach2* (whenever it detects spikes at all) appear to perform rather well. Which model is to be preferred based on the CDR and FDR jointly is a matter of the goal of the forecaster. In particular, it strongly depends how costly it is not to detect a spike and to forecast a spike when in fact none occurs. Christensen et al. (2012) argued that not forecasting a spikes is in fact more costly than a false detection. This argumentation can be backed by the fact, that price spikes have the potential to yield prices up to \$12 500.

While it is indeed important to be able to forecast the occurrence of spikes, it may very often be sufficient to have a good forecast for the probability of a spike. We denote the forecasted spike probability as \hat{h}_t and, as before, a spike realization as S_t . Then the out-of-sample forecast performance can be evaluated by empirical loss functions obtained from T^* out-of-sample observations. In the following, we use the mean absolute error (MAE) defined as

$$\text{MAE} = \frac{1}{T^*} \sum_{t=1}^{T^*} |S_t - \hat{h}_t| \quad (4.1)$$

and the root mean square error (RMSE) given by

$$\text{RMSE} = \sqrt{\frac{1}{T^*} \sum_{t=1}^{T^*} (S_t - \hat{h}_t)^2}. \quad (4.2)$$

Since the consequences of falsely detecting a spike are different from falsely not detecting an extreme price, Christensen et al. (2012) suggested to measure the forecast performance additionally by an asymmetric loss function. Following their approach, we therefore consider also the asymmetric mean absolute error (AMAE) given by

$$\text{AMAE} = \frac{1}{T^*} \sum_{t=1}^{T^*} (1 - \kappa + 2 \kappa S_t) |S_t - \hat{h}_t|. \quad (4.3)$$

Table 5: Spike detections

	VIC		NSW		QLD		SA	
Threshold	A\$100	A\$300	A\$100	A\$300	A\$100	A\$300	A\$100	A\$300
Window I								
Spikes	451	16	299	29	267	10	450	8
Detections								
<i>ach</i> ₁	218	0	205	16	66	0	0	0
<i>ach</i> ₂	244	2	244	0	153	0	233	0
<i>ach</i> ₃	296	0	232	0	183	0	232	0
<i>logit</i> ₁	451	16	299	29	267	10	450	0
<i>logit</i> ₂	215	0	235	15	154	3	213	0
<i>logit</i> ₃	356	5	265	21	209	2	383	0
<i>logit</i> ₄	425	9	305	35	265	3	450	2
CDR								
<i>ach</i> ₁	35.03	0.00	51.84	34.48	21.72	0.00	0.00	0.00
<i>ach</i> ₂	52.33	0.00	74.92	0.00	53.18	0.00	50.67	0.00
<i>ach</i> ₃	60.75	0.00	62.54	0.00	58.05	0.00	57.33	0.00
<i>logit</i> ₁	82.93	50.00	82.61	62.07	77.15	40.00	86.22	0.00
<i>logit</i> ₂	45.68	0.00	72.24	34.48	53.18	20.00	46.22	0.00
<i>logit</i> ₃	70.73	31.25	77.59	51.72	68.54	20.00	77.11	0.00
<i>logit</i> ₄	79.16	43.75	84.62	65.52	77.53	20.00	86.44	12.50
FDR								
<i>ach</i> ₁	27.52	NaN	24.39	37.50	12.12	NaN	NaN	NaN
<i>ach</i> ₂	3.28	NaN	8.20	NaN	7.19	NaN	2.15	NaN
<i>ach</i> ₃	6.80	NaN	10.53	NaN	16.22	NaN	3.73	NaN
<i>logit</i> ₁	17.07	50.00	17.39	37.93	22.85	60.00	13.78	NaN
<i>logit</i> ₂	4.19	NaN	8.09	33.33	7.79	33.33	2.35	NaN
<i>logit</i> ₃	10.39	0.00	12.45	28.57	12.44	0.00	9.40	NaN
<i>logit</i> ₄	16.00	22.22	17.05	45.71	21.89	33.33	13.56	50.00
Window II								
Spikes	55	11	201	38	184	37	107	29
Detections								
<i>ach</i> ₁	0	0	148	92	392	6	47	7
<i>ach</i> ₂	30	0	79	48	67	0	15	0
<i>ach</i> ₃	0	0	35	0	67	0	0	0
<i>logit</i> ₁	55	11	201	38	184	37	107	29
<i>logit</i> ₂	35	9	156	74	140	30	67	12
<i>logit</i> ₃	36	10	174	62	158	31	68	12
<i>logit</i> ₄	46	11	187	77	174	30	104	19
CDR								
<i>ach</i> ₁	0.00	0.00	58.71	76.32	31.52	0.00	22.43	17.24
<i>ach</i> ₂	45.45	0.00	35.82	52.63	29.89	0.00	11.21	0.00
<i>ach</i> ₃	0.00	0.00	14.93	0.00	29.35	0.00	0.00	0.00
<i>logit</i> ₁	70.91	63.64	85.57	76.32	75.54	54.05	56.07	55.17
<i>logit</i> ₂	49.09	54.55	71.14	92.11	61.96	54.05	44.86	34.48
<i>logit</i> ₃	49.09	63.64	77.61	92.11	71.20	54.05	45.79	34.48
<i>logit</i> ₄	56.36	63.64	82.09	94.74	73.37	54.05	55.14	41.38
FDR								
<i>ach</i> ₁	NaN	NaN	20.27	68.48	85.20	100.00	48.94	28.57
<i>ach</i> ₂	16.67	NaN	8.86	58.33	17.91	NaN	20.00	NaN
<i>ach</i> ₃	NaN	NaN	14.29	NaN	19.40	NaN	NaN	NaN
<i>logit</i> ₁	29.09	36.36	14.43	23.68	24.46	45.95	43.93	44.83
<i>logit</i> ₂	22.86	33.33	8.33	52.70	18.57	33.33	28.36	16.67
<i>logit</i> ₃	25.00	30.00	10.34	43.55	17.09	35.48	27.94	16.67
<i>logit</i> ₄	32.61	36.36	11.76	53.25	22.41	33.33	43.27	36.84

Note: This table reports the correct and false spike detections for thresholds equal to A\$100/MWh and A\$300/MWh for the out-of-sample forecasts for 1 July to 30 September 2007 (window I) and 1 January to 31 December 2011 (window II). CDR and FDR are reported in percent. A NaN is reported whenever no spikes were predicted and the corresponding rate could not be determined.

The asymmetric mean absolute error is a special case of the generalized loss function proposed by Elliott et al. (2005) and gives more weight to falsely not forecasted spikes than to false detections. Using the value $\kappa = 0.5$ suggested by Christensen et al. (2012) failing to forecast an actual spike is penalized three times more than a false detection. Based on these three criteria we again compute the MCS. Due to the shorter out-of-sample period the block length for the bootstrap is chosen to correspond to one week of data. The results can be found in Table 6. Note that for the best performing model the loss is highlighted with an asterisk, whereas it is reported in bold whenever the corresponding model belongs to the MCS. The results are mixed, but a few observations can be made. First, more models belong to the MCS than for the in-sample period, which can be explained by the shorter out-of-sample period resulting in less powerful hypothesis tests. Second, no single model stands out as the best performing one. In fact, each model shows the best performance for a specific market, threshold and time window. However, the *ach3*, *logit3* and *logit4* specifications rank first and are included more often than any other model. Furthermore, the *ach1* model shows the overall worst performance by being performing best or being in the MCS least often. The results continue to be mixed when focusing on a specific loss function, threshold choice and time window, although some differences in performance can be observed.

Based solely on these results the *ach3*, *logit3* and *logit4* models appear to be the most recommendable ones. Combining the results from Table 6 with the ones in Tables 5 and 4 one can conclude that the *logit3* and the *logit4* specification might be the most promising ones. They show good in-sample performance and their forecasts strike a good balance between correct and false detections.

5. Conclusion

Accurately forecasting price spikes is essential for all market participants. Retailers could better hedge their positions while producers could include the information into their bidding strategies. Even industrial consumers that are capable of switching energy consumption over short periods of time could potentially profit from accurate spike forecasts. The importance of forecasting spikes has already been pointed out by Christensen et al. (2012), who proposed a sophisticated model to achieve that goal.

In the paper at hand we reconsider this problem by first providing a detailed analysis of the data under consideration. Most importantly we show that spikes do often occur in blocks and that they exhibit a certain autocorrelation structure. Based on these characteristics we suggest extensions to the model proposed by Christensen et al. (2012) and alternatives in form of logit specifications which are inspired by the proposals of Kauppi and Saikkonen (2008). These specifications are shown to have a superior fit when evaluated both in-sample and out-of-sample.

Instead of solely comparing the models in terms of certain loss functions we compute the model confidence set (MCS) proposed by Hansen et al. (2011) in order to identify models that are equivalent in their performance. Two dynamic logit models denoted *logit3* and *logit4* that allow for regime switching behavior by distinguishing blocks of spikes show the best overall performance. These models take the fact into account that once a spike has occurred it is very likely to observe another spike. As exogenous variables we included temperatures and forecasted electricity loads, which help to capture the seasonalities at different frequencies. Although following Christensen et al. (2012) we solely consider half hourly one-step forecasts, practitioners could easily adapt the presented models to forecast spike probabilities over horizons longer than 30 minutes.

For future research it may be useful to study and forecast the multivariate behavior of price spikes. Such information may be valuable for e.g. buyers and sellers who are active on more than one market and try to minimize their overall price risk. In this regard first contributions for daily data and on the complete price distribution have already been made by, e.g., Worthington et al. (2005), Higgs (2009) and Ignatieva and Trueck (2011).

Table 6: Model confidence set

	VIC		NSW		QLD		SA	
Threshold	A\$100	A\$300	A\$100	A\$300	A\$100	A\$300	A\$100	A\$300
Window I: RMSE								
<i>ach</i> ₁	23.25	5.80	18.08	6.90	19.73	4.71	29.87	4.23
<i>ach</i> ₂	18.52	5.81	13.13*	7.53	15.27	4.74	17.74	3.93*
<i>ach</i> ₃	17.71	5.76	14.97	7.82	16.22	4.75	17.26	4.21
<i>logit</i> ₁	18.07	5.25	14.67	6.38	15.61	4.43	16.70	4.21
<i>logit</i> ₂	19.35	4.97	13.70	6.13	15.22	4.29	18.90	3.99
<i>logit</i> ₃	17.32	4.68*	13.19	5.87*	14.23*	4.23*	16.18	3.94
<i>logit</i> ₄	16.98*	4.87	13.54	6.19	14.51	4.65	15.46*	4.02
Window I: MAE								
<i>ach</i> ₁	11.43	0.56	7.11	1.13	7.47	0.47	10.46	0.31*
<i>ach</i> ₂	6.95	0.61	4.38	1.03	5.19	0.45	7.09	0.54
<i>ach</i> ₃	6.48	0.58	4.86	0.90	5.72	0.36*	6.74	0.37
<i>logit</i> ₁	5.66*	0.47*	3.79*	0.71*	4.08*	0.46	5.50*	0.39
<i>logit</i> ₂	7.88	0.69	5.27	1.10	5.70	0.74	7.98	0.51
<i>logit</i> ₃	6.30	0.61	4.57	0.97	4.66	0.73	6.06	0.52
<i>logit</i> ₄	6.61	0.64	4.83	1.02	5.19	0.83	6.22	0.55
Window I: AMAE								
<i>ach</i> ₁	9.05	0.29*	5.57	0.74	5.26	0.24	5.87	0.16*
<i>ach</i> ₂	5.37	0.33	3.37*	0.62	3.88	0.23	5.62	0.30
<i>ach</i> ₃	5.15*	0.32	3.69	0.49*	4.55	0.18*	5.44	0.19
<i>logit</i> ₁	5.43	0.35	3.58	0.56	3.70	0.29	5.26	0.24
<i>logit</i> ₂	6.30	0.43	4.26	0.80	4.33	0.45	6.44	0.28
<i>logit</i> ₃	5.24	0.39	3.73	0.72	3.63*	0.44	5.18*	0.29
<i>logit</i> ₄	6.00	0.44	4.29	0.83	4.39	0.54	5.78	0.33
Window II: RMSE								
<i>ach</i> ₁	5.62	2.24	7.64	6.07	13.33	4.71	7.36	3.77
<i>ach</i> ₂	4.10	2.01	7.07	4.33	7.86	4.36	6.81	3.94
<i>ach</i> ₃	5.44	2.50	8.39	4.50	7.87	4.54	7.41	3.99
<i>logit</i> ₁	3.97	1.93	5.61	3.02*	6.72	3.87	6.63	3.40
<i>logit</i> ₂	3.95*	1.87	5.33	4.08	6.57	3.65*	6.17	3.11*
<i>logit</i> ₃	3.96	1.80*	5.19	3.68	6.34	3.69	6.13*	3.15
<i>logit</i> ₄	3.96	1.84	5.11*	3.83	6.24*	3.69	6.29	3.23
Window II: MAE								
<i>ach</i> ₁	1.07	0.18	1.77	0.73	3.06	0.53	1.65	0.31
<i>ach</i> ₂	0.31	0.09	0.86	0.30	1.01	0.33	0.78	0.23
<i>ach</i> ₃	0.41	0.08*	1.02	0.26*	1.03	0.28*	0.81	0.22
<i>logit</i> ₁	0.67	0.15	0.97	0.27	1.18	0.38	1.19	0.31
<i>logit</i> ₂	0.33	0.08*	0.60	0.32	0.96	0.32	0.81	0.20*
<i>logit</i> ₃	0.38	0.08*	0.59	0.29	0.89	0.32	0.84	0.20*
<i>logit</i> ₄	0.33*	0.08*	0.57*	0.30	0.84*	0.32	0.77*	0.20*
Window II: AMAE								
<i>ach</i> ₁	0.57	0.10	1.17	0.72	2.86	0.31	1.01	0.18
<i>ach</i> ₂	0.20*	0.06	0.66	0.27	0.71	0.18	0.48	0.12
<i>ach</i> ₃	0.21	0.04*	0.72	0.14*	0.72	0.14*	0.44*	0.11*
<i>logit</i> ₁	0.43	0.09	0.76	0.20	0.89	0.25	0.82	0.20
<i>logit</i> ₂	0.24	0.06	0.49	0.33	0.74	0.23	0.56	0.12
<i>logit</i> ₃	0.26	0.06	0.48*	0.27	0.67*	0.23	0.57	0.12
<i>logit</i> ₄	0.26	0.06	0.52	0.30	0.69	0.24	0.59	0.14

Note: This table reports the RMSE, MAE and AMAE losses (multiplied by 100) for spike forecasts with thresholds (thr) equal to \$100/MWh and \$300/MWh for the out-of-sample forecasts for July 1 to September 30 2007 (window I) and January 1 to December 31 2011 (window II). The loss for the best performing model is marked with an asterisk. For models belonging to the model confidence set (MCS) the loss is reported in bold. To get the MCS we used block-bootstrap with block-length of 336 and 5 000 simulations. The tests were based on a significance level of 5%.

References

- AEMO, 2010. An Introduction to Australia’s National Electricity Market. Australian Energy Market Operator, Melbourne.
- Anderson, E. J., Hu, X., Winchester, D., 2007. Forward contracts in electricity markets: the Australian experience. *Energy Policy* 35, 3089–3103.
- Chan, K. F., Gray, P., Van Campen, B., 2008. A new approach to characterizing and forecasting electricity price volatility. *International Journal of Forecasting* 28, 728–743.
- Christensen, T., Hurn, S., Lindsay, K., 2009. It never rains but it pours: modeling the persistence of spikes in electricity prices. *The Energy Journal* 30, 25–48.
- Christensen, T., Hurn, S., Lindsay, K., 2012. Forecasting spikes in electricity prices. *International Journal of Forecasting* 24, 400–411.
- Elliott, G., Komunjer, I., Timmermann, A., 2005. Estimation and testing of forecast rationality under flexible loss. *Review of Economic Studies* 72, 1107–1125.
- Engle, R. F., Russel, J. R., 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica* 66, 1127–1162.
- Fan, S., Hyndman, R. J., 2012. Short-term load forecasting based on a semi-parametric additive model. *IEEE Transactions on Power Systems* 27, 134–141.
- Fernandes, M., Grammig, J., 2006. A family of autoregressive conditional duration models. *Journal of Econometrics* 130, 1–22.
- Hamilton, J. D., Jorda, O., 2002. A model of the federal funds rate target. *Journal of Political Economy* 110, 1135–1167.
- Hansen, P. R., Lunde, A., Nason, J. M., 2011. The model confidence set. *Econometrica* 79, 453–497.
- Higgs, H., 2009. Modelling price and volatility inter-relationships in the Australian wholesale spot electricity markets. *Energy Economics* 31, 748–756.
- Ignatieva, K., Trueck, S., 2011. Modeling spot price dependence in Australian electricity markets with applications to risk management. Working paper, Macquarie University Sydney.
- Janczura, J., Weron, R., 2010. An empirical comparison of alternate regime-switching models for electricity spot prices. *Energy Economics* 32, 1059–1073.
- Kanamura, T., Ohashi, K., 2008. On transition probabilities of regime switching in electricity price. *Energy Economics* 30, 1158–1172.
- Kauppi, H., Saikkonen, P., 2008. Predicting U.S. recessions with dynamic binary response models. *The Review of Economics and Statistics* 90, 777–791.
- Knittel, C. R., Roberts, M. R., 2005. An empirical examination of restructured electricity prices. *Energy Economics* 27, 792–817.
- Mount, T. D., Ning, Y., Cai, X., 2006. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics* 28, 62–80.
- Rambharat, R. B., Brockwell, A. E., Seppi, D. J., 2005. A threshold autoregressive model for wholesale electricity prices. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 54, 287–299.
- Taylor, J. W., 2003. Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society* 54, 799–805.

- Taylor, J. W., 2010. Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles. *International Journal of Forecasting* 26, 627–646.
- Taylor, J. W., 2012. Short-term load forecasting with exponentially weighted methods. *IEEE Transactions on Power Systems* 27, 458–464.
- Taylor, J. W., de Menezes, L. M., McSharry, P. E., 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting* 22, 1–16.
- Thomas, S., Ramiah, V., Mitchell, H., Heaney, R., 2011. Seasonal factors and outlier effects in rate of return on electricity spot prices in Australia’s national electricity market. *Applied Economics* 43, 355–369.
- Weron, R., 2006. Modeling and forecasting electricity loads and prices: a statistical approach. Wiley, Chichester.
- Worthington, A., Kay-Spratley, A., Higgs, H., 2005. Transmission of prices and price volatility in Australian electricity spot markets: a multivariate GARCH analysis. *Energy Economics* 27, 337–350.